DOCUMENT RESUME

ED 222 506 TM 820 638

AUTHOR Natriello, Gary
TITLE The Impact of the Evaluation of Teaching on Teacher Effort and Effectiveness.
SPONS AGENCY Spencer Foundation, Chicago, Ill.
PUB DATE Mar 82
NOTE 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Elementary School Teachers; *Evaluation Methods; Intermediate Grades; Measurement Objectives; Personnel Evaluation; Self Evaluation (Individuals); Surveys; *Teacher Effectiveness; *Teacher Evaluation; *Teacher Improvement; Teacher Participation; *Teaching Skills
IDENTIFIERS *Teacher Attitudes Toward Evaluation; Teacher Surveys

ABSTRACT

The effects of performance evaluations of teachers on their effort and effectiveness was studied. The evaluation system based on Dornbusch and Scott's theory was used. Stages in evaluating performance included allocating, criteria-setting, sampling, appraising, communicating the evaluation, and planning for improvement. Sets of tasks to characterize teachers' work included teaching subject matter, character development, maintaining control, and record keeping. The author developed a concept termed "leverage" which referred to the relationship between subordinate effort and subordinate outcomes. Teachers in six inner-city schools responded to Likert Scale response surveys dealing with the evaluation of their performance, the effort they devoted to teaching, and their effectiveness at those tasks. Two measures of leverage were obtained, Teacher Assessment of Leverage (TAL) and Composite Assessment of Leverage (CAL). Frequency of evaluation was positively related to teacher leverage. Teachers evaluated more frequently experienced a greater degree of effectiveness in relation to their effort on teaching tasks. It was found that tasks were conceptualized as inert by teachers. Confirmation of this finding would carry negative consequences for learning experiences of students and teachers' work experiences. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED222506

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

G. Natriello

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

The Impact of the Evaluation of Teaching on Teacher Effort and Effectiveness

Gary Natriello
Washington University

TM 820 635

Paper Presented at the Annual Meeting of the American Educational Research Association, New York, March 1982

The Research Reported Here Was Made Possible by a Grant from the Spencer Foundation.

The Impact of the Evaluation of Teaching on Teacher Effort and Effectiveness

During the last decade there has been increased pressure for teacher accountability and a great deal of interest in the evaluation of teachers. (Natriello, et al., 1977) Most discussions of the evaluation of teaching treat it as a means for achieving greater control over the educational process. An overriding concern in such work is that schools are not performing as efficiently or effectively as they should. Nevertheless, the work on the evaluation of teaching, although dealing at length with methods and procedures for assessing teacher performance and student outcomes, (Millman, 1981) pays little attention to the impact of such evaluations on teachers or students. Thus we have very little insight as to how procedures used to evaluate teaching affect the efficiency or effectiveness of teachers in their execution of their teaching tasks.

This paper examines the effects of the evaluation of teachers on their effort and effectiveness in the conduct of their teaching tasks. Unlike most inquiries into the evaluation of teachers, the current study investigates the impact of evaluation from the perspective of teachers themselves. It thus departs from the dominant rational model of school organizations which assumes that decisions and policies instituted by administrative superiors are implemented by those further down in the organization.

Theory

The current study was guided by the theory of evaluation and the exercise of authority developed by Dornbusch and Scott through studies of evaluation practices in a variety of organizations. (Dornbusch and Scott, 1975) A series of studies conducted by Dornbusch and Scott and their colleagues revealed that there were benefits to be derived from increasing the frequency of the evaluation of teachers in schools. In these studies (For a summary of these studies see Natriello and Dornbusch, 1982.) teachers reported greater satisfaction and less inclination to complain of problems in the evaluation system when evaluation was more frequent. Furthermore, the evaluation of teachers appeared to be a rather infrequent event in most public schools. (Natriello and Rowe, 1981).

The theory developed by Dornbusch and Scott has several features which make it particularly useful for the present investigation of the evaluation of teaching. First, their theory is based on an explicit model of the evaluation process. They specify six stages or types of activities that make up a complete evaluation system. The figure below depicts the six stages in a complete model of the system for evaluating performance.

Figure 1 About Here

Allocating refers to the process of assigning a task to an individual performer, in the present case, a teacher. Before a task can be evaluated, it must have been assigned. Obviously, if no one is given an assignment to complete a task, there will be no one and no performance to evaluate.

After a task has been allocated performers may know how they are supposed to perform the task, but may not know what dimensions of the task are important. In general, supervisors are interested in more dimensions of a task than simply whether an attempt was made to perform it. Therefore, in order to make assessments of performance along desired dimensions, evaluators must specify criteria against which the performance of the task is to be assessed. Criteria setting is the second stage in the model of the evaluation process. Dornbusch and Scott (1975, pp. 138) note that criteria setting involves, first, determining which properties of the task should be considered in the assessment of the performance; second, deciding the relative weight of each property in the overall evluation; and third, setting a standard for performance along each dimension.

Sampling is the third stage of the evaluation model. Sampling refers to the process of collecting information on the performance of the allocated task according to the criteria for performance that have been set. Sampling involves two separate decisions about what information to collect on task performance. The first decision is over the choice of indicators for determining the performance. The second decision is the selection of the sampling technique for gathering the information.

The fourth stage in the evaluation model is appraisal. Appraisal

is the act of assigning an evaluation to a performance. The appraisal stage entails bringing together the criteria set for the task performance with the sample of information collected on that performance to arrive at an evaluation. To the extent that the procedures for the criteria setting and sampling stages of the process have been spelled out, the appraisal stage will be more straightforward. Nevertheless, there are always decisions which must be made by evaluators at this stage. Evaluators must make inferences from the sample of information collected on performance to the full performance. Often the supervisor must rely on past experience and rules of thumb to make such decisions. Evaluators must also decide how the standards set for the performance are to be applied to a specific performance. This involves a fair amount of evaluator discretion since evaluators must judge how comparable the specific performance situation is to the situations for which the standards were developed.

Appraisal is an activity that takes place in the mind of the evaluator. Obviously, if evaluators expect those they are evaluating to react to their evaluations, the evaluations must be communicated. The fifth stage of the evaluation model involves the communication of the results of the evaluation by the evaluator to the subordinate.

Most evaluations turn up at least one area where performance is not as good as might be desirable. When unsatisfactory evaluations are communicated to those being evaluated, it is often the case that simply communicating dissatisfaction to those being evaluated will not result in desired improvement. Performers may not know how to improve

performance or they may lack the resources to do so. Thus it is important for evaluators and performers to engage in planning for improvement. This may involve changes in performer behavior as well as changes in the arrangements set up by the evaluator for the performer.

The six stage model of the process of performance evaluation is obviously arbitrary in its division of activities. Nevertheless, it represents fairly well the scope and sequence of activities involved in the evaluation of individual performance in organizations. At times these activities are not performed explicitly. At other times these activities are not performed at all. Many of the tasks expected of performers in organizations are never included as part of the formal system for performance evaluation. Even considering both formal and informal evaluation, teachers may receive very infrequent evaluation as Dornbusch and Scott found when they studied elementary schools. The model provides us with a formal analysis of the procedures that would be involved in a performance evaluation system. It thus guides our inquiry into the evaluation of teachers in schools.

A second feature of the the theory developed by Dornbusch and Scott that makes it particularly appropriate for our present work is its task-specific nature. Much of the research on work in organizations (e.g. Woodward, 1970) rests on what Dornbusch and Scott (1975, p. 77) label an "assumption of homogeneity." That is, it assumes that the technology used by an organization is similar across tasks and positions and that the hierarchical relations of the organization are uniform across work groups. The theory of evaluation and the exercise of authority instead focuses on the variations in

technologies and work arrangements associated with different tasks performed by individuals in a single position in the organization. It thus permits us to assemble a more sensitive portrait of the work of teachers in schools.

In their studies of the evaluation of the performance of teachers Dornbusch and Scott developed the following set of tasks to characterize the work of teachers:

> Teaching Subject Matter: including, for example, leading discussions; preparing lesson plans; lecturing; examining and grading students on their knowledge of subject matter.
>
> Character development: including, for example, social skills, morality; interpersonal relations.
>
> Maintaining Control: including, for example, preventing interferences with other classes; keeping down the noise level; keeping the attention of the class on their work.
>
> Record Keeping: including, for example, reporting absences, and/or latecomers; maintaining administrative records; turning in grades accurately and on time; taking attendance.

They found that the technologies used by teachers varied across such tasks as teaching subject matter and record keeping and that the supervisory arrangements varied also. In the present study we adopted these teaching tasks as a way to conceptualize the work of teachers.

Dornbusch and Scott developed the notion of "task conceptions" to refer to participants' beliefs concerning the characteristics of their tasks. They distinguished two polar types of these task conceptions, inert and active. They note that "A task is considered inert to the extent that participants believe that resistance to carrying out task activities is predictable from performance to performance. A task is considered active to the extent that participants believe that the resistance to carrying out task activities is unpredictable from

performance to performance." (Dornbusch and Scott, 1975, p. 348) They demonstrate that there is a relationship between task conceptions held by performers and preferences for supervisory and work arrangements. In the present study we employ this distinction between active and inert task conceptions in discussing the relationship between evaluation practices and teacher effort and effectiveness.

Armed with the model of the evaluation process, a task-specific focus, and the concept of task conceptions, all part of the theory of evaluation and the exercise of authority, we set out to develop hypotheses to guide our examination of the relationship between the evaluation of teaching and teacher effort and effectiveness. The earlier work by Dornbusch and Scott and their colleagues had demonstrated a positive relationship between frequency of evaluation and teacher satisfaction with the evaluation system. As we attempted to extend the theory to deal with more than feelings of teacher satisfaction, we reasoned that frequency of evaluation was a useful feature of the evaluation system to examine. Not only did the earlier work on the evaluation of teachers demonstrate the positive relationship between frequency of evaluation and teacher satisfaction with the evaluation system (Natriello and Dornbusch, 1982), but research on organizations has long examined the relationship between closeness of supervision and employee reactions. (Perrow, 1972) The strong positive relation between frequency of evaluation and satisfaction for teachers should not, as Dornbusch and Scott point out, be interpreted as demand for very frequent evaluations, but rather as dissatisfaction with the infrequent evaluations typically found in public schools. (Dornbusch and Scott, 1975, p. 168) They

thus suggest the possibility of a curvilinear relationship between frequency of evaluation and teacher satisfaction, with too little or too much evaluation promoting dissatisfaction.

Moving beyond teacher satisfaction to examine the effects of evaluation systems on concepts more closely related to teacher performance leads us to a serious conceptual problem. We might on the one hand follow the lead of some earlier investigators who have examined the impact of supervisory and evaluative procedures on subordinate effort. However, this approach is open to the criticism that it ignores the final outcomes of that effort. On the other hand, we might follow those who have studied the impact of supervisory and evaluative procedures on the outcomes of subordinate performance. This approach is, however, open to the critique that it fails to take into account the differences in task difficulty or the variation in the predictability of accomplishing different tasks. The first approach would seem to lead to the development of evaluation systems designed to maximize subordinate effort with little regard for the payoff from those efforts, while the second would seem to lead to the development of evaluation systems that would dissuade organizations and individuals from taking on new and unpredictable tasks, tasks that present greater risk of failure even when greater effort is devoted to them.

To overcome the weaknesses of these two approaches we developed a concept which takes into account both subordinate effort and subordinate outcomes. This new concept, which we term "leverage", refers to the relationship between the effort put forth by a subordinate and the outcomes which result from that effort. Leverage

involves the relationship between a worker and his or her work, more specifically, between the effort exerted by a worker and the outcome or product of that effort on his or her tasks. Leverage may be described as the ratio between outcome and effort and may range from 1/0 to 0/ 1. 1/0 means that for 0 units of effort the individual realizes 1 unit of outcome. 0/1 means that the individual realizes 0 units of outcome for every 1 unit of effort. Of course, 1/1 means that for every unit of effort the individual realizes 1 unit of outcome.

The use of the concept of leverage casts the investigation of the impact of evaluation on subordinates in a new light. Heretofore, investigations of the impact of managerial strategies, including evaluation, on subordinates have been concerned with increasing the control over workers by examining the impact on their effort or with increasing control over the work by examining the impact on outputs. In either case the concern was with increasing managerial control with little or no attention devoted to subordinates. The use of the concept of leverage shifts the focus to a concentration on enabling subordinates to gain control over their work and thus only indirectly improving organizational or managerial control. Strategies designed to increase leverage should improve the relationship between individual workers and their work.

The above discussion suggests that it would be fruitful to develop greater understanding of the relationship between the evaluation practices used in schools and teacher leverage. Following along the lines of Dornbusch and Scott who suggested a curvilinear relationship between frequency of evaluation and teacher satisfaction,

we proposed the following hypothesis:

> There is a curvilinear relationship between frequency of evaluation and teacher leverage. The more frequent the evaluation of teaching, the greater will be the leverage of teachers up to the point of too close supervision.

A second hypothesis speaks to the differences in task conceptions as they effect the relationship between frequency of evaluation and leverage:

> There is a curvilinear relationship between task conceptions and the impact of evaluation on teacher leverage. The more active the task conception, the greater the impact of evaluation on leverage up to the point of a task being too active for supervisors to make useful contributions.

This second hypothesis is based in part on the argument that in the case of very inert task conceptions evaluation would have little impact on leverage since the task is taken to be so straightforward at the outset. In addition, in the case of very active task conceptions evaluation would have little impact on leverage since the task resistance is so unpredictable that supervisors may be able to make a very small or no contribution to improving the ratio between effectiveness and effort. Evaluation should thus have the greatest positive impact on teacher leverage when it occurs with just the right frequency in the case of tasks that are moderate in terms of the active-inert dimension.

Method

The study was conducted in six inner-city middle schools of a large midwestern city. In May 1981 teachers in these six schools were asked to complete a survey dealing with the evaluation of teaching by their principal. Of a total population of 189 classroom teachers, 182 completed the survey, for a response rate of 97%. We enjoyed this unusually high response rate primarily because members of our research team had spent one day a week for the previous five months working in each school. Results of the observations are reported elsewhere. (deCharms and Natriello, 1981) Data for the present paper come from the teacher survey.

The survey contained items used in the contruction of the major indices for the present analysis. Teachers were asked to describe the frequency with which their principals engaged in five of the six stages of the model of the evaluation process outlined by Dornbusch and Scott: task assignment, criteria-setting, sampling, providing feedback, and planning for improvement. No items were used concerning the frequency of appraisal since the appraisal process goes on in the mind of the principal and is not observable by teachers. For each of the remaining five stages of the evaluation process we had one or more items which asked teachers to describe the frequency with which they experienced the relevant activities. Teachers responded in terms of Likert scale response categories. These five subscales were combined into a scale of frequency of evaluation. Since teachers responded to questions about these five stages in terms of the four teaching tasks, we have a scale of frequency of evaluation for each teaching task.

Three summary measures of frequency of evaluation were

constructed. The first measure provided an indication of the presence of frequent activity on any of the six stages of the evaluation process. It was constructed by counting each of the five subscales as a 1 if the teacher response indicated either frequent or very frequent activity for that stage of the evaluation process. Teachers were divided into two groups, those who reported no stage of the process having frequent or very frequent activity, and those reporting one or more stages with frequent or very frequent activity.

A second measure indicating the number of stages with frequent activity was constructed along the same lines. This time teachers were divided into three groups, those reporting no stage with frequent activity, those reporting one stage with frequent activity, and those reporting two or more stages with frequent actvity.

A third summary measure was used to indicate the relative frequency of evaluation activity. The third measure was based on two submeasures. The first submeasure was a simple additive index of teacher responses for questions on the five stages. The second submeasure ranked teachers according to their highest (most frequent) response to any of the five stages of the evaluation process. To construct our measure of relative frequency of evaluation these two submeasures were combined. For each submeasure, teachers were divided into three groups corresponding to high, medium, and low frequency of evaluation. Cases falling into the same group on both measures were included in the measure of relative frequency. Those falling into inconsistent groups were excluded from the analyses involving this summary measure.

Two measures of leverage were obtained through the teacher

surveys. The first, called the Teacher Assessment of Leverage (TAL), was derived from an item on the survey which asked teachers to describe the relationship between their effort and the payoff to that effort during the past year of teaching. Teachers selected from among nine response categories composed of one of three effort ratings (high,medium,low) and one of three payoff ratings (high, medium, low). On the basis of these responses teachers were classified into two groups, those for whom the payoff was less than the effort (leverage < 1), and those for whom the payoff was equal to or greater than their effort (leverage $\geq$ 1). The Teacher Assessment of Leverage measure was obtained for each of the teaching tasks.

A second leverage measure, termed the Composite Assessment of Leverage (CAL) was constructed from two other items on the teacher survey. One item asked teachers to rate their current effectiveness in comparison with their effectiveness earlier in their teaching career. The second item asked teachers to rate their current effort in comparision with their effort earlier in their teaching career. In both cases teachers indicated their rating in terms of one of five response categories (For effectiveness: much less effective, less effective, about as effective, more effective, much more effective; For effort, described as how hard you are working: not nearly as hard, not quite as hard, about as hard, somewhat harder, and much harder). The Composite Assessment of Leverage was constructed by taking the ratio of effectiveness to effort and placing teachers into one of two groups, those whose effectiveness was less than their effort (leverage < 1) and those whose effectiveness was equal to or greater than their effort (leverage $\geq$ 1) The Composite Assessment of

Leverage measure was created for each of the four teaching tasks.

A third set of summary measures were constructed to reflect the task conceptions held by teachers of their teaching tasks. Following the procedures used by Dornbusch and Scott, three items on the teacher survey asked teachers to rate the goal clarity, predictability,. and efficacy of each of their teaching tasks. In each case teacher responses were collapsed into three categories reflecting high, medium, or low ratings of these aspects of the active-inert dimension. An additive scale was constructed using these three items for each of the four teaching tasks.

Results

The results of analyses to examine the first hypothesis are presented in Tables 1, 2, and 3.

Insert Tables 1,2, and 3 About Here

Table 1 reports the relationship between the first summary measure of frequency of evaluation (the presence of frequent activity in a least one stage of the evaluation process) and the two measures of teacher leverage (self-assessment of leverage and composite assessment of leverage). The gammas for the relationships are positive for all four teaching tasks for both measures of leverage. In most cases the

differences in the proportions of systems in the two conditions resulting in leverage greater than or equal to 1 are substantial. Only in the case of self-assessment of leverage for the record keeping task and composite assessment of leverage for the task of teaching subject matter are the relationships between the presence of frequent evaluation activity and leverage weak.

Table 2 reports a parallel set of analyses using the second summary measure of frequency of evaluation, the number of evaluation stages with frequent activity. Once again all of the gammas are positive and most of the relationships are quite strong. Only in the cases of self-assessment of leverage for the record keeping task and composite assessment of leverage for the task of teaching subject matter are the relationships between the number of evaluation stages with frequent activity and teacher leverage weak.

Table 3 reports the results of the analyses using the third summary measure of frequency of evaluation, the relative frequency of evaluation. Again all of the gammas are positive and rather strong, the weakest relationship involving self-assessment of leverage for the task of record keeping (gamma = .28).

The first hypothesis of a curvilinear relationship between frequency of evaluation and teacher leverage was not confirmed. Although there is considerable support for the argument that increasing the frequency of evaluation leads to increased teacher leverage, no evidence was found to suggest that after a certain point continuing to increase the frequency of evaluation leads to diminished teacher leverage.

The composite assessment of leverage measure permitted an

analysis of the effects of frequency of evaluation on teacher effectiveness and teacher effort separately. Since our leverage measure is a ratio of effectiveness to effort, leverage could increase if the greater frequency of evaluation either led to enhanced effectiveness or diminished effort. While the first would be desirable, clearly the second would not. Table 4 reports the results of the analyses designed to determine whether the increase in teacher leverage associated with more frequent evaluation is due to increased effectiveness or decreased effort.

Insert Table 4 About Here

Section A of Table 4 reveals that the relationship between the presence of frequent evaluation activity and teacher effectiveness is positive and quite strong for each of the four teaching tasks. The gammas range from .51 to .63. Section B of Table 4 shows that the relationships between the presence of frequent evaluation activity and teacher effort are in general positive and weak. The one exception is for the task of maintaining control where the gamma is -.03. Increasing the frequency of evaluation leads to greater teacher effectiveness, not to diminished teacher effort on teaching tasks.

It was not possible to examine our second hypothesis with the teachers in our present sample since very few of the teachers held

active task conceptions for their teaching tasks. When teacher responses to the three items from the survey tapping the aspects of the active-inert dimension are trichotomized into high, medium, and low (high being active, low being inert) only 9% of the teachers in the current sample report the task of teaching subject matter to be highly active on one of these dimensions. The corresponding figures are 9%, 9.5%, and 5.8% for the tasks of character development, maintaining control, and record keeping. Even by this very wide interpretation of active task conceptions a very small proportion of the teachers display such active task conceptions. The teachers conceive of all of their teaching tasks as quite inert.

Discussion

The present study did not confirm the original hypothesis of a curvilinear relationship between frequency of evaluation and teacher leverage. Frequency of evaluation was positively related to leverage at all levels. Furthermore, the enhancement of teacher leverage results primarily from an increase in reported effectiveness in the conduct of teaching tasks. Put another way, increasing the frequency with which teachers are evaluated seems to lead to teachers working smarter, not harder. Under conditions of more frequent evaluation teachers report being markedly more effective with the same or slightly more effort. Although there was a strong and consistent positive relationship between the frequency of evaluation and teacher leverage, the present study provided no evidence to suggest that after a point increasing the frequency of evaluation leads to diminished teacher leverage. This may be due to the fact that the evaluation of

teachers in public schools is seldom too frequent. Dornbusch and Scott (1975) noted that the evaluation of teachers was quite infrequent as did Natriello and Rowe (1981). The survey data and additional observations from the six schools in the present study suggest that here too the evaluation of teachers was generally an infrequent event.

The second hypothesis of a curvilinear relationship between task conceptions and the impact of evaluation on teacher leverage could not be examined because there was not sufficient variation in teacher task conceptions. In their earlier studies Dornbusch and Scott found that teachers conceive of three of the four teaching tasks as active, that is, as low in clarity, predictability, and efficacy. In contrast, the teachers in the six inner-city schools in the present study conceived of all of their teaching tasks are inert, that is, as high in clarity, predictability, and efficacy. These differences between the findings of Dornbusch and Scott and the findings of the present study may be due to a number of other differences in the studies. Dornbusch and Scott collected their data in the early 1970's, while data for the present study was collected in 1981. The passage of time might have led to changes in the conceptions of teaching tasks. The data for the earlier studies came from teachers in middle-class suburban schools on the San Francisco penninsula, while the data for the present study came from teachers in schools serving lower class predominantly black students in the inner city of a major metropolitan center in the midwest. Teachers in the earlier studies worked in small to medium size districts, while teachers in the present study worked in a large urban district.

These factors suggest that the differences in findings might be explained by the increasing proletarianization of teaching. Braverman (1974) has observed that in advanced capitalist societies white collar work becomes increasingly bureaucratized and proletarianized, i.e. the tasks become routinized. Bowles and Gintis (1975) have made the same observation. Wise (1979) has discussed this process as it effects schools under the rubric of hyperrationalization and has identified recent changes in educational policy that are likely to lead to the bureaucratization of the classroom. The general phenomenon of the proletarianization of teaching might explain our finding that teachers in a large urban system in 1981 conceive of their teaching tasks as more inert than teachers in smaller systems in 1970.

Conclusion

The failure to confirm the original hypotheses should not obscure the two rather dramatic findings. First, the present study suggests that teacher leverage can be increased through more frequent evaluations of teaching. Administrators and policy makers interested in enhancing teacher perceptions of effectiveness and quite possibly actual effectiveness might consider evaluation as a tool to accomplish this goal. Those who oppose the use of teacher evaluation because they perceive it as a control mechanism injurious to the professional autonomy of teachers might consider the positive effects of such practices in enhancing the control of teachers over their teaching tasks.

Second, the finding that teachers in the inner-city schools in the study held very inert conceptions of their teaching tasks might be more closely examined as evidence of the increasing proletarianization

of teaching. The finding that teaching tasks are increasingly conceptualized as inert by teachers, if confirmed, would carry serious negative consequences not only for the learning experiences of students but also for the work experiences of teachers.

Figures and Tables

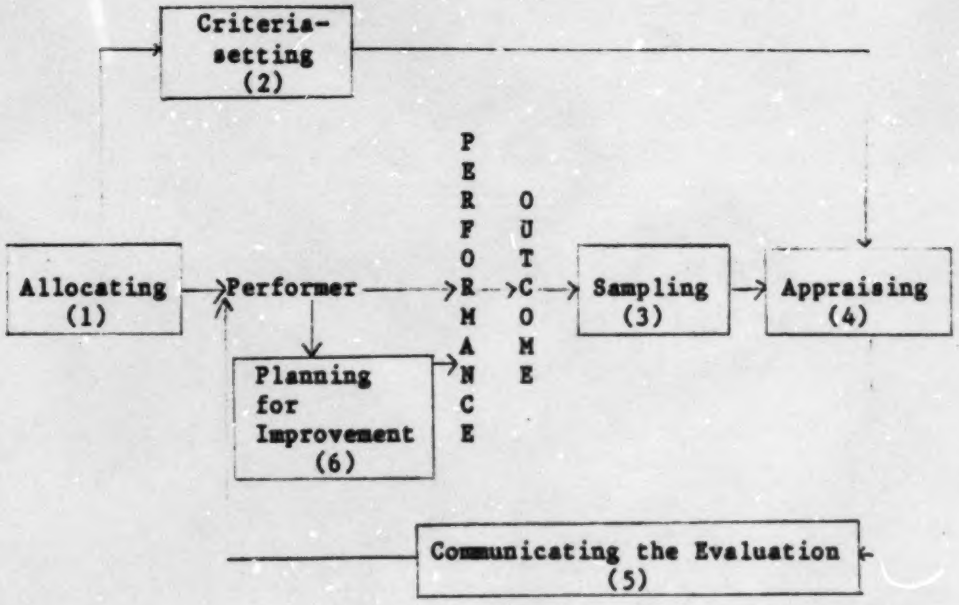

Figure 1 A Model of the Evaluation Process

Table 1

Relation of Frequent Evaluation Activity and Teacher Leverage for Four Teaching Tasks

	Gamma	Proportion of Evaluation Systems With at Least One Stage Showing Frequent Activity Resulting in Leverage ≥ 1	Proportion of Evaluation Systems With No Stage Showing Frequent Activity Resulting in Leverage ≥ 1
A. Self-Assessment of Leverage (SAL)			
Teaching Tasks			
Teaching Subject Matter	.61	.59	.26
Character Development	.35	.61	.43
Maintaining Control	.52	.61	.32
Record Keeping	.08	.79	.77
B. Composite Assessment of Leverage (CAL)			
Teaching Tasks			
Teaching Subject Matter	.13	.58	.51
Character Development	.51	.71	.44
Maintaining Control	.35	.63	.44
Record Keeping	.49	.84	.64

Table 2

Relation of Number of Evaluation Stages with Frequent Activity and Teacher Leverage for Four Teaching Tasks

	Gamma	Proportion of Evaluation Systems with 2 or More Stages Showing Frequent Activity Resulting in Leverage $\geq$ 1	Proportion of Evaluation Systems with No Stage Showing Frequent Activity Resulting in Leverage $\geq$ 1
A. Self-Assessment of Leverage (SAL)			
Teaching Tasks			
Teaching Subject Matter	.61	.71	.26
Character Development	.41	.73	.43
Maintaining Control	.51	.69	.32
Record Keeping	.14	.83	.77
B. Composite Assessment of Leverage (CAL)			
Teaching Tasks			
Teaching Subject Matter	.22	.68	.51
Character Development	.50	.79	.44
Maintaining Control	.41	.73	.44
Record Keeping	.43	.86	.64

Table 3

Relation of Relative Frequency of Evaluation and Teacher Leverage for Four Teaching Tasks

	Gamma	Proportion of Evaluation Systems Showing High Relative Frequency of Activity Resulting in Leverage $\geq$ 1	Proportion of Evaluation Systems Showing Low Relative Frequency of Activity Resulting in Leverage $\geq$ 1
A. Self-Assessment of Leverage (SAL)			
Teaching Tasks			
Teaching Subject Matter	.62	.77	.23
Character Development	.32	.75	.45
Maintaining Control	.60	.80	.35
Record Keeping	.28	.86	.72
B. Composite Assessment of Leverage (CAL)			
Teaching Tasks			
Teaching Subject Matter	.33	.79	.50
Character Development	.51	.89	.46
Maintaining Control	.44	.78	.47
Record Keeping	.40	.84	.61

Table 4

Relation of Frequent Evaluation Activity and Teacher Effort and Effectiveness for Four Teaching Tasks

A. Effectiveness

Teaching Tasks	Gamma	Proportion of Evaluation Systems With at Least One Stage Showing Frequent Activity Resulting in Greater Effectiveness	Proportion of Evaluation Systems With No Stage Showing Frequent Activity Resulting in Greater Effectiveness
Teaching Subject Matter	.51	.69	.38
Character Development	.59	.69	.33
Maintaining Control	.53	.63	.32
Record Keeping	.63	.62	.26

B. Effort

Teaching Tasks	Gamma	Proportion of Evaluation Systems With at Least One Stage Showing Frequent Activity Resulting in Greater Effort	Proportion of Evaluation Systems With No Stage Showing Frequent Activity Resulting in Greater Effort
Teaching Subject Matter	.24	.77	.64
Character Development	.21	.71	.59
Maintaining Control	-.03	.67	.67
Record Keeping	.06	.37	.32

References

Bowles, S. and H. Gintis. SCHOOLING IN CAPITALIST AMERICA. New York: Basic Books, Inc., Publishers, 1976.

Braverman, H. LABOR AND MONOPOLY CAPITAL: THE DEGREDATION OF WORK IN THE TWENTIETH CENTURY. New York: Monthly Review Press, 1974.

deCharms, R. and G. Natriello. "Evaluation and Teacher Motivation." St. Louis: Report to the Spencer Foundation, 1981.

Dornbusch, S.M. and W.R. Scott. EVALUATION AND THE EXERCISE OF AUTHORITY. San Francisco: Jossey-Bass Publishers, 1975.

Millman, Jason, ed. HANDBOOK OF TEACHER EVALUATION. Beverly Hills: Sage Publications, 1981.

Natriello, G., T. Deal, S.M. Dornbusch, and M. Hoag. "A Summary of the Recent Literature on the Evaluation of Principals, Teachers, and Students." Stanford, California: Stanford Center for Research and Development in Teaching, 1977.

Natriello, G. and B. Rowe. "Life in a Loosely-Coupled World: The Evaluation of Elementary School Teachers." St. Louis: Washington University, Unpublished Paper, 1981.

Natriello, G. and S.M. Dornbusch. "Pitfalls in the Evaluation of Teachers by Principals." ADMINISTRATOR'S NOTEBOOK, forthcoming.

Perrow, C. COMPLEX ORGANIZATIONS: A CRITICAL ESSAY. Glenview, Illinois: Scott Foresman, 1972.

Wise, A. LEGISLATED LEARNING: THE BUREAUCRATIZATION OF THE AMERICAN CLASSROOM. Berkeley: University of California Press, 1979.

Woodward, J. INDUSTRIAL ORGANIZATION: BEHAVIOR AND Control. London: Oxford University Press, 1970.